Datasets in the IPYNB file:

|  |  |
| --- | --- |
| **Name** | **Description** |
| Census\_1 | All the raw data together |
| Population\_df | First population variables |
| Poverty\_df | Key poverty variables |
| Housing\_df | Key housing variables |
| Edu\_df | Key education variables |
| Lang\_df | Key language variables |
| Zip\_Geo\_df | Geo data for the zipcode |
| Zip\_Geo\_klaar | Zip\_Geo data with states |
| Basic\_Pop | All Zipgeo data and population data together, and pop density |
| Edu \_scores | Education scores for each zipcode, with Zipcode and Pop |
| Rur\_urb\_df | Has rural/urban type for every zipcode |
| Full\_pop\_stats | Has Basic\_Pop and Rur\_urb\_df together |
| Testing | Just for performing tests |
| Final\_SE\_df | This is the final table where all final numbers will go. |
| \_ToMerge | Each of Poverty\_df, Housing\_df, and Lang\_df have ToMerge files, with columns removed that wouldn’t go into the main df |
| Hospitals\_df | All the hospitals, with all the indicators, from Connor |
| Zipcodes\_df | When we read in Final\_SE\_df to work with hospitals |
| Group \_HZ | The number of hospitals in each zipcode |
| Zips\_heat\_num\_hos | CSV created, paired down to track num hospitals per zip |
| Hospitals\_df\_paired | Hospitals\_df with many columns taken out and renamed |
| Zip\_PD\_ToMerge | Paired down with just Zip and Pop Density to merge with Hospitals\_df\_paired |
| Hosp\_PD | All the hospital data, but now with the Pop Density for the zipcode where the hosp is located |
|  |  |

Consider removing all zipcodes with 0 population, and all zipcodes for certain variables where they have a perfect score or 0 score.